PATENT APPLICATION/

ATTORNEY DOCKET NO. ___10007912-1___

(stamp: OIPE JC59 JUL 22 2004)

IN THE
UNITED STATES PATENT AND TRADEMARK OFFICE

Inventor(s): Maria Castellanos et al.	**Confirmation No.:** 3894
Application No.: 09/944,919	**Examiner:** Wayne P. Amsbury
Filing Date: August 31, 2001	**Group Art Unit:** 2171
Title: METHOD AND SYSTEM FOR MINING A DOCUMENT CONTAINING DIRTY TEXT	

Mail Stop Appeal Brief-Patents
Commissioner For Patents
PO Box 1450
Alexandria, VA 22313-1450

TRANSMITTAL OF APPEAL BRIEF

Sir:

Transmitted herewith in **triplicate** is the Appeal Brief in this application with respect to the Notice of Appeal filed on __May 19, 2004__ .

07/22/2004 STEUMEL1 00000025 082025 09944919
The fee for filing this Appeal Brief is (37 CFR 1.17(c)) $330.00. 01 FC:1402 330.00 DA

(complete (a) or (b) as applicable)

The proceedings herein are for a patent application and the provisions of 37 CFR 1.136(a) apply.

( ) (a) Applicant petitions for an extension of time under 37 CFR 1.136 (fees: 37 CFR 1.17(a)-(d)
for the total number of months checked below:

( )	one month	$110.00
( )	two months	$420.00
( )	three months	$950.00
( )	four months	$1480.00

( ) The extension fee has already been filled in this application.

(X) (b) Applicant believes that no extension of time is required. However, this conditional petition is being made to provide for the possibility that applicant has inadvertently overlooked the need for a petition and fee for extension of time.

Please charge to Deposit Account **08-2025** the sum of ___$330.00___. At any time during the pendency of this application, please charge any fees required or credit any over payment to Deposit Account 08-2025 pursuant to 37 CFR 1.25. Additionally please charge any fees to Deposit Account 08-2025 under 37 CFR 1.16 through 1.21 inclusive, and any other sections in Title 37 of the Code of Federal Regulations that may regulate fees. A duplicate copy of this sheet is enclosed.

(X) I hereby certify that this correspondence is being deposited with the United States Postal Service as first class mail in an envelope addressed to: Commissioner for Patents, Alexandria, VA 22313-1450. Date of Deposit:__July 19, 2004__
OR
( ) I hereby certify that this paper is being transmitted to the Patent and Trademark Office facsimile number_____ on _____

Number of pages:

Typed Name: Steven E. Dicke

Signature: *Steven E. Dicke*

Rev 05/04 (Aplbrief)

Respectfully submitted,
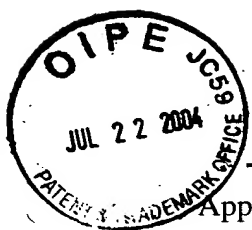
Maria Castellanos et al.

By *Steven E. Dicke*

Steven E. Dicke
Attorney/Agent for Applicant(s)
Reg. No. **38,431**

Date: **July 19, 2004**

Telephone No.: **(612) 573-2002**

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Appellant: Maria Castellanos et al.	Examiner: Wayne P. Amsbury
Serial No.: 09/944,919	Group Art Unit: 2171
Filed: August 31, 2001	Docket No.: 10007912-1
Title: METHOD AND SYSTEM FOR MINING A DOCUMENT CONTAINING DIRTY TEXT	

APPEAL BRIEF TO THE BOARD OF PATENT APPEALS AND INTERFERENCES OF THE UNITED STATES PATENT AND TRADEMARK OFFICE

Mail Stop Appeal Brief-Patents
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Dear Sir/Madam:

Appellant's Brief on Appeal

This brief is presented in support of the Notice of Appeal filed on May 21, 2004, from the final rejection dated March 4, 2004, and the Advisory Action dated June 29, 2004, of the Office Action rejecting claims 1-12 and 15-35 of the above identified application. Thirty-three claims remain for consideration.

The Appeal Brief is filed in triplicate. The U.S. Patent and Trademark Office is hereby authorized the Charge Deposit Account No. 08-2025 in the amount of $330.00 for filing a Brief in Support of an Appeal as set forth under 37 C.F.R. 1.17(c), however, at any time during the pendency of this application, please charge any fees required or credit any overpayment to Deposit Account 08-2025 pursuant to 37 C.F.R. 1.25. Additionally, please charge any fees to Deposit Account 08-2025 under 37 C.F.R. 1.16, 1.17, 1.19, 1.20 and 1.21.

Appellant respectfully requests reversal of the Office Action's rejection of pending claims 1-12 and 15-35.

07/22/2004 STEUMEL1 00000025 09944919
01 FC:1402   330.00 DA

**Appeal Brief to the Board of Patent Appeals and Interferences
of the United States Patent and Trademark Office**
Appellant: Maria Castellanos et al.
Serial No.: 09/944,919
Filed: August 31, 2001
Docket No.: 10007912-1
Title: METHOD AND SYSTEM FOR MINING A DOCUMENT CONTAINING DIRTY TEXT

Table of Contents

Appeal Brief to the Board of Patent Appeals and Interferences
of the United States Patent and Trademark Office
Appellant: Maria Castellanos et al.
Serial No.: 09/944,919
Filed: August 31, 2001
Docket No.: 10007912-1
Title: METHOD AND SYSTEM FOR MINING A DOCUMENT CONTAINING DIRTY TEXT

Real Party in Interest

The real party in interest is Hewlett-Packard Company, LP.

Related Appeals and Interferences

There are no other appeals or interferences known to Appellants which will have a bearing on the Board's decision in the present Appeal.

Status of the Claims

Claims 1-12 and 15-35 are pending, and are the subject of the present Appeal (see Appendix A).

Claims 1-12 and 15-35 are rejected under 35 U.S.C. § 103(a) as being unpatentable over U.S. Patent No. 6,085,206 to Domini et al. (Domini) in view of U.S. Patent No. 4,965,763 to Zamora (Zamora). No claims have been allowed. Claims 13 and 14 have been cancelled. Claims 1-12 and 15-35 are appealed herein.

Status of the Amendments

A Response was filed on April 27, 2004 subsequent to the Final Office Action mailed March 4, 2004. The claims listed in Appendix A reflect the claims as of March 4, 2004.

Summary of the Invention

The present invention, as claimed in independent claim 1, provides a computer-implemented method for mining a document containing dirty text. The method includes removing an instance of dirty text within the document to produce a clean document having a content. A data mining operation is performed on the cleaned document thereby deriving relevant information from the cleaned document and providing a summary of the content of the document. See Figs. 1-3, pages 6-16.

In another embodiment, the present invention, as claimed in independent claim 11, provides a computer system including a bus, a memory unit coupled to the bus, and a processor coupled to the bus, the processor for executing a method for mining a document containing dirty text. The method includes producing a clean document having a content

Appeal Brief to the Board of Patent Appeals and Interferences
of the United States Patent and Trademark Office
Appellant: Maria Castellanos et al.
Serial No.: 09/944,919
Filed: August 31, 2001
Docket No.: 10007912-1
Title: METHOD AND SYSTEM FOR MINING A DOCUMENT CONTAINING DIRTY TEXT

comprising performing a general cleaning of the document by removing an instance of dirty text within the document including instances of misspelling and grammatical errors, and performing a domain and task specific cleaning of the document including removing instances of computer code and tables to produce a clean document. A data mining operation is performed on the cleaned document including providing a summary of the content of the document. See Figs. 1-3, pages 6-16.

In another embodiment, the present invention, as claimed in independent claim 21, provides a computer-usable medium having computer-readable program code embodied therein for causing a computer system to remove an instance of dirty text within the document to produce a clean document having a content. A data mining operation is performed on the clean document to provide a summary of the content. See Figs. 1-3, pages 6-16.

In another embodiment, the present invention as claimed in independent claim 31, provides a computer-implemented method for mining a document containing dirty text. The method includes producing a clean document having a content comprising performing a general cleaning of said document by removing one or more instance of dirty text within the document including instances of misspelling and grammatical errors, and performing a domain and task specific cleaning of the document including removing instances of computer codes and tables. A data mining operation is performed on the clean document, including determining a sentence score for each sentence of the clean document and ranking the sentences from highest to lowest based on the sentence score. A summary of the content of the document is generated using the highest rank sentences. See Figs. 1-3, pages 6-16.

Issues Presented for Review

Whether the rejection of claims 1-12 and 15-35 in the Final Office Action mailed under 35 U.S.C. § 103(a) as being unpatentable over Domini in view of Zamora, sets forth a case of *prima facie* obviousness.

Appeal Brief to the Board of Patent Appeals and Interferences
of the United States Patent and Trademark Office
Appellant: Maria Castellanos et al.
Serial No.: 09/944,919
Filed: August 31, 2001
Docket No.: 10007912-1
Title: METHOD AND SYSTEM FOR MINING A DOCUMENT CONTAINING DIRTY TEXT

Grouping of the Claims

The claims do not stand or fall together, but are grouped as follows and each group is believed to be patentable.

A. Claims 1-10 and 21-30, with claim 1 being representative of the group.

B. Claims 11, 12 and 15-20.

C. Claims 31-34.

D. Claim 35.

Each group is separately patentable. Arguments for separate patentability are given below in the Argument Sections A – D for each respective group.

Argument

A. The Rejection of Claims 1-10 under 35 U.S.C. § 103(a) based on Domini in view of Zamora

The rejection of claims 1-10 under 35 U.S.C. § 103(a) as being unpatentable over Domini in view of Zamora fails to set forth a *prima facie* case of obviousness and should be withdrawn. Appellant submits that Domini, either alone or in view of Zamora, fails to teach or suggest the invention of independent claim 1.

Referring to Section 706.02 (j) of the M.P.E.P., to establish a *prima facie* case of obviousness, three basic criteria must be met:

(1) There must be some suggestion or motivation, either in the references themselves or in the knowledge generally available to one of ordinary skill in the art, to modify the reference of to combine reference teachings;

(2) There must be a reasonable expectation of success;

(3) The prior art reference (or references when combined) must teach or suggest all the claim limitations.

The teaching or suggestion to make the claimed combination and the reasonable expectation of success must both be found in the prior art and not based on Appellant's disclosure. See In re Vaeck, 947 F.2d 488, 20 U.S.P.Q.2d 1438 (FED. Cir. 1991).

Appeal Brief to the Board of Patent Appeals and Interferences
of the United States Patent and Trademark Office
Appellant: Maria Castellanos et al.
Serial No.: 09/944,919
Filed: August 31, 2001
Docket No.: 10007912-1
Title: METHOD AND SYSTEM FOR MINING A DOCUMENT CONTAINING DIRTY TEXT

Independent claim 1 recites a computer-implemented method for mining a document containing dirty text. The method includes removing an instance of dirty text within the document to produce a cleaned document having a content. The method also includes performing a data mining operation on the cleaned document thereby deriving relevant information from the cleaned document and providing a summary of the content of the document.

Domini is directed to identifying dirty text in a document and provides both spell and grammar checking in a document at the same time. Domini specifically defines dirty text as that text which has not been spell checked and/or that has not been grammar checked (See Domini, column 9, lines 43-48). Furthermore, Domini describes that after a sentence has been grammar checked, it is marked as clean text (column 9, lines 49-53).

Zamora merely recites a computer method for automatic extraction of commonly specified information from business correspondence. The method utilizes a parametric information extraction (PIE) system to identify fields of a business document such as frame slots for a business correspondence or list of business correspondence closing phrases (See Zamora, Fig. 3 and Fig. 5). The identified fields disclosed are limited to "standardized forms" (Col. 3, l. 36) such as "author, date, recipient, address, subject statement . . ." (Col. 3, ll. 23-24).

Domini fails to disclose **performing a data mining operation on a cleaned document.** The Office Action cites Domini (Col. 13, lines 19-42 and Col. 4, lines 10-29) for teaching or suggesting this claimed limitation. Appellant respectfully disagrees.

Domini (at Col. 13, lines 19-42) teaches an AutoCorrect button wherein "every time that the user types the misspelled word 315 in the document (or in any other document until the user deletes the AutoCorrect entry) the misspelled word will be automatically changed to the suggestion 320 selected by the user from the suggestion list box 317." Appellant contends that automatically corrected misspelled words in a document is not "performing a data mining operation."

Domini (at Col. 4, lines 10-29) teaches a method of spell checking and grammar checking. For convenience, this section of Domini is reproduced below:

Appeal Brief to the Board of Patent Appeals and Interferences
of the United States Patent and Trademark Office
Appellant: Maria Castellanos et al.
Serial No.: 09/944,919
Filed: August 31, 2001
Docket No.: 10007912-1
Title: METHOD AND SYSTEM FOR MINING A DOCUMENT CONTAINING DIRTY TEXT

> More particularly described, the present invention provides a method for spell checking and grammar checking a document. A sentence is parsed from the document. It is determined whether any of the words in the sentence are misspelled and an indication, such as presenting the misspelled word in red, bold typeface, is provided for any misspelled words. In response, the user can then provide an input command that is indicative of the changes to be made to any misspelled words, such as ignore, change, etc. These steps are repeated until all of the misspelled words in the sentence have been indicated to the user.

> It is then determined whether the sentence that was parsed from the document is grammatically proper. If not, an indication is provided to designate the portion of the sentence that is improper. For instance, the improper word or words may be displayed to the user in green, bold typeface. The user, in response, can provide an input command that indicates any changes for the sentence or document. Each grammatically improper portion of the sentence can be separately displayed.

Domini, thus, teaches parsing a sentence from the document, determining if any spelling or grammar errors occur in the parsed sentence, and correcting (via user input) the errors. For several reasons, Appellants contend that this section of Domini does not teach or suggest the claimed limitiation.

First, parsing a sentence from a document and determining if errors exist in the sentence is not "performing a data mining operation." Second, Domini teaches and suggests a specific sequence of steps. Namely, a sentence is first parsed from the document. Then, a determination is made if any errors exist in the sentence. If errors exist, then an indication of such errors is presented to the user. Compare, though, this teaching with the claimed limitation: performing a data mining operation **on a clean document**. Domini does not teach or suggest performing data mining operations on a clean document. By contrast, Domini clearly teaches away from this claim limitation, as expressed in the following passage:

> For example, after a sentence has been spell checked and grammar checked, it is marked with a "clean" spell check flag and a "clean" grammar check flag. The flags indicate that the **text does not need to be checked again** by the spell and grammar check functions. It is possible for text to be "clean" for

Appeal Brief to the Board of Patent Appeals and Interferences
of the United States Patent and Trademark Office
Appellant: Maria Castellanos et al.
Serial No.: 09/944,919
Filed: August 31, 2001
Docket No.: 10007912-1
Title: METHOD AND SYSTEM FOR MINING A DOCUMENT CONTAINING DIRTY TEXT

> spell checking and "dirty" for grammar checking, and vice
> versa. After text has been marked "clean" for spelling, then the
> spell checker program module is able to **skip over** this text
> when spell checking. Similarly, when a range of text has been
> marked "clean" for grammar checking, then the grammar
> checker program module is able to **skip over** this text when
> grammar checking. Because "clean" text does not need to be
> checked, the speed of the spell checker program module and
> grammar checker program module is increased for the
> examination of a previously checked document. (Col. 9, lines
> 49-64: Emphasis Added)

Domini and Zamora, alone or in combination, fail to teach or suggest other claimed limitations as well. The references fail to disclose performing a data mining operation on the clean document **thereby deriving relevant information from said clean document and providing a summary of the content of said document.** In fact, the Office Action even concedes that "Domini does not explicity address providing a summary of content ..." (Paper No. 6, page 2). Zamora, alone or in combination with Domini, fails to cure this deficiency.

The Office Action contends that the "index of Zamora corresponds to a summary, so does the inverted file of FIG. 18" (Paper No. 6, page 3). Appellant respectfully disagrees. Zamora merely uses a parametric information extraction system to identify fields of a business document, such as author, dates, recipient, address, etc., and does not disclose providing a summary of the content of a cleaned document.

Further, Zamora fails to disclose removing an instance of dirty text within said document to produce a cleaned document having a content. In fact, the Office Action admits: "Zamora...does not explicitly correct spelling and other forms of dirty text..." (Paper No. 6, page 2). Since neither Domini nor Zamora teach or suggest performing a data mining operation on a cleaned document thereby deriving relevant information from said cleaned document and providing a summary of the content of said document, one skilled in the art could not apply the teachings of Domini in view of Zamora and arrive at the present invention of independent claim 1. In fact, Zamora teaches away from cleaning a document prior to performing a data mining operation, since Zamora is triggering on specific and expected business letter fields like closing phrases and headers.

Appeal Brief to the Board of Patent Appeals and Interferences
of the United States Patent and Trademark Office
Appellant: Maria Castellanos et al.
Serial No.: 09/944,919
Filed: August 31, 2001
Docket No.: 10007912-1
Title: METHOD AND SYSTEM FOR MINING A DOCUMENT CONTAINING DIRTY TEXT

Accordingly, Appellant respectfully requests that the above rejection under 35 U.S.C. § 103(a) should be withdrawn. Dependant claims 2-10 depend directly or indirectly upon independent claim 1. Accordingly, dependant claims 2-10 are also allowable over the art of record.

B. The Rejection of Claims 11, 12 and 15-20 under 35 U.S.C. § 103(a) based on Domini in view of Zamora

The rejection of claims 11, 12 and 15-20 under 35 U.S.C. § 103(a) as being unpatentable over Domini in view of Zamora fails to set forth a prima facie case of obviousness and should be withdrawn. Appellant submits that Domini, either alone or in view of Zamora, fails to teach or suggest the invention of independent claim 11.

Referring to Section 706.02 (j) of the M.P.E.P., to establish a *prima facie* case of obviousness, three basic criteria must be met:

(1) There must be some suggestion or motivation, either in the references themselves or in the knowledge generally available to one of ordinary skill in the art, to modify the reference of to combine reference teachings;
(2) There must be a reasonable expectation of success;
(3) The prior art reference (or references when combined) must teach or suggest all the claim limitations.

The teaching or suggestion to make the claimed combination and the reasonable expectation of success must both be found in the prior art and not based on Appellant's disclosure. See In re Vaeck, 947 F.2d 488, 20 U.S.P.Q.2d 1438 (FED. Cir. 1991).

Claim 11 recites a computer system. The computer system includes a bus, a memory unit coupled to the bus, and a processor coupled to the bus. The processor executes a method for mining a document containing dirty text. The method includes producing a cleaned document having a content including performing a general cleaning of the document by removing an instance of dirty text within the document including instances of misspelling and grammatical errors, performing a domain and task-specific cleaning of the document including removing instances of computer code and tables to produce a cleaned document. A data mining operation is performed on the cleaned document including providing a summary of the content of the document.

Appeal Brief to the Board of Patent Appeals and Interferences
of the United States Patent and Trademark Office
Appellant: Maria Castellanos et al.
Serial No.: 09/944,919
Filed: August 31, 2001
Docket No.: 10007912-1
Title: METHOD AND SYSTEM FOR MINING A DOCUMENT CONTAINING DIRTY TEXT

For at least the reasons as stated above in Section A with reference to independent claim 1, Appellant believes independent claim 11 to be allowable over Domini either alone or in view of Zamora.

Additionally, neither Domini nor Zamora teaches or suggests the claimed cleaning process. Specifically, neither Domini nor Zamora teaches or suggests **performing a general cleaning of said document by removing an instance of dirty text within said document including instances of misspelling and grammatical errors, and performing a domain and task specific cleaning of said document including removing instances of computer codes and tables.** In fact, the Office Action concedes the following: "neither Domini nor Zamora explicitly addresses the extraction of computer code or a table," then merely states "but these are well known components of documents of various types" (Paper No. 6, page 3).

In light of the admission in the Office Action and lack of teachings and suggestions in the references, one skilled in the art could not combine the teachings of Domini in view of Zamora and arrive at the present invention of independent claim 11.

Dependent claims 12, and 15-20 depend either directly or indirectly upon independent claim 11. Accordingly, these dependent claims are allowable over the art of record.

C. The Rejection of Claims 31-34 under 35 U.S.C. § 103(a) based on Domini in view of Zamora

The rejection of claims 31-34 under 35 U.S.C. § 103(a) as being unpatentable over Domini in view of Zamora fails to set forth a *prima facie* case of obviousness and should be withdrawn. Appellant submits that Domini, either alone or in view of Zamora, fails to teach or suggest the invention of independent claim 31.

Referring to Section 706.02 (j) of the M.P.E.P., to establish a *prima facie* case of obviousness, three basic criteria must be met:

> (1)   There must be some suggestion or motivation, either in the references themselves or in the knowledge generally available to one of ordinary skill in the art, to modify the reference of to combine reference teachings;
> (2)   There must be a reasonable expectation of success;
> (3)   The prior art reference (or references when combined) must teach or suggest all the claim limitations.

Appeal Brief to the Board of Patent Appeals and Interferences
of the United States Patent and Trademark Office
Appellant: Maria Castellanos et al.
Serial No.: 09/944,919
Filed: August 31, 2001
Docket No.: 10007912-1
Title: METHOD AND SYSTEM FOR MINING A DOCUMENT CONTAINING DIRTY TEXT

The teaching or suggestion to make the claimed combination and the reasonable expectation of success must both be found in the prior art and not based on Appellant's disclosure. See In re Vaeck, 947 F.2d 488, 20 U.S.P.Q.2d 1438 (FED. Cir. 1991).

Independent claim 31 recites a computer-implemented method for mining a document containing dirty text. The method includes producing a cleaned document having a content comprising performing a general cleaning of said document by removing one or more instance of dirty text within said document including instances of misspelling and grammatical errors, and performing a domain and task specific cleaning of said document including removing instances of computer codes and tables. A data mining operation is performed on said cleaned document, including determining a sentence score for each sentence of said cleaned document and ranking the sentences from highest to lowest based on the sentence score. A summary of the content of the document is generated using the highest ranked sentences.

For at least the reasons as stated above in Section A with reference to independent claim 1, Appellant believes independent claim 31 to be allowable over Domini either alone or in view of Zamora.

Additionally, nothing in the art of record teaches or suggests **determining a sentence score for each sentence of said cleaned document and ranking the sentences from highest to lowest based on the sentence score to provide a summary based on the highest ranked sentences,** after completion of the claimed cleaning process. To address this limitation, the Office Action states: "ranking is implicitly from highest-to-lowest in one direction and lowest-to-highest in the other' (Paper No. 6, page 4).

Further, the Office Action references a scoring in Zamora (Col 2, ll. 24-31), but Zamora is limited to determining how many occurrences there are of a user-defined search term in various documents that are being searched and then ranking the various documents. The Office Action concedes that Zamora does not rank individual sentences, but still attempts to apply the reference. Again, Zamora fails to disclose determining a sentence score for each sentence of said cleaned document and ranking the sentences as recited in independent claim 31. In view of the above, one skilled in the art could not combine the teachings of Domini in view of Zamora and arrive at the present invention of independent claim 31.

Appeal Brief to the Board of Patent Appeals and Interferences
of the United States Patent and Trademark Office
Appellant: Maria Castellanos et al.
Serial No.: 09/944,919
Filed: August 31, 2001
Docket No.: 10007912-1
Title: METHOD AND SYSTEM FOR MINING A DOCUMENT CONTAINING DIRTY TEXT

Accordingly, Appellant respectfully requests that the above rejection of independent claim 31 under 35 U.S.C. § 103(a) should be withdrawn. Dependant claims 32-34 depend directly or indirectly upon independent claim 31, they are also allowable over the art of record.

D. The Rejection of Claim 35 under 35 U.S.C. § 103(a) based on Domini in view of Zamora

The rejection of claim 1-35 under 35 U.S.C. § 103(a) as being unpatentable over Domini in view of Zamora fails to set forth a *prima facie* case of obviousness and should be withdrawn. Appellant submits that Domini, either alone or in view of Zamora, fails to teach or suggest the invention of independent claim 35.

Referring to Section 706.02 (j) of the M.P.E.P., to establish a *prima facie* case of obviousness, three basic criteria must be met:

> (1)   There must be some suggestion or motivation, either in the references themselves or in the knowledge generally available to one of ordinary skill in the art, to modify the reference of to combine reference teachings;
> (2)   There must be a reasonable expectation of success;
> (3)   The prior art reference (or references when combined) must teach or suggest all the claim limitations.

The teaching or suggestion to make the claimed combination and the reasonable expectation of success must both be found in the prior art and not based on Appellant's disclosure. See In re Vaeck, 947 F.2d 488, 20 U.S.P.Q.2d 1438 (FED. Cir. 1991).

For at least the reasons as stated above in Section C with reference to independent claim 31, Appellant believes dependent claim 35 to be allowable over Domini either alone or in view of Zamora.

Additionally neither Domini nor Zamora teach or suggest **determining a sentence score for each sentence including applying a keyword technique to each sentence (claim 32); applying a location technique to each sentence (claim 33); and applying a semantic similarity technique to each sentence (claim 34); wherein the semantic similarity technique comprises generating a vector associated with each sentence, and comparing each vector to every other vector, including defining a co-sign of an angle between two vectors and using the co-sign of the angle between two vectors to determine whether**

Appeal Brief to the Board of Patent Appeals and Interferences
of the United States Patent and Trademark Office
Appellant: Maria Castellanos et al.
Serial No.: 09/944,919
Filed: August 31, 2001
Docket No.: 10007912-1
Title: METHOD AND SYSTEM FOR MINING A DOCUMENT CONTAINING DIRTY TEXT

sentences represented by the two vectors are semantically related (claim 35). None of these claim elements are taught or suggested by Domini, either alone or in view of Zamora. One skilled in the art could not combine the teachings of Domini in view of Zamora and arrive at the invention of claim 35. Accordingly, Appellant respectfully requests that the above rejection of claim 35 under 35 U.S.C. § 103(a) be withdrawn.

Conclusion

For above reasons, Appellants respectfully submit that the cited art fails to render the claimed invention obvious. Therefore, Appellants respectfully submit that the rejections to pending claims 1-12 and 15-25 are in error. Appellants respectfully request that the Board reverse the Office Action and find all pending claims allowable.

The U.S. Patent and Trademark Office is hereby authorized the Charge Deposit Account No. 08-2025 in the amount of $330.00 for filing a Brief in Support of an Appeal as set forth under 37 C.F.R. 1.17(c), however, at any time during the pendency of this application, please charge any fees required or credit any overpayment to Deposit Account 08-2025 pursuant to 37 C.F.R. 1.25. Additionally, please charge any fees to Deposit Account 08-2025 under 37 C.F.R. 1.16, 1.17, 1.19, 1.20 and 1.21.

Any inquiry regarding this Appeal Brief to the Board of Patent Appeals and Interferences of the United States Patent and Trademark Office should be directed to either Steven E. Dicke at Telephone No. (612) 573-2002, Facsimile No. (612) 573-2005 or Howard Boyle at Telephone No. (281) 514-9645, Facsimile No. (281) 514-8332. In addition, all correspondence should continue to be directed to the following address:

Hewlett-Packard Company
Intellectual Property Administration
P.O. Box 272400
3404 E. Harmony Road, M/S 35
Fort Collins, Colorado 80527-2400

Appeal Brief to the Board of Patent Appeals and Interferences
of the United States Patent and Trademark Office
Appellant: Maria Castellanos et al.
Serial No.: 09/944,919
Filed: August 31, 2001
Docket No.: 10007912-1
Title: METHOD AND SYSTEM FOR MINING A DOCUMENT CONTAINING DIRTY TEXT

Respectfully submitted,

Maria Castellanos et al.,

By their attorneys,

DICKE, BILLIG & CZAJA, PLLC
Fifth Street Towers, Suite 2250
100 South Fifth Street
Minneapolis, MN 55402
Telephone: (612) 573-2002
Facsimile: (612) 573-2005

Date: July 19, 2004

SED:jan

Steven E. Dicke
Reg. No. 38,431

CERTIFICATE UNDER 37 C.F.R. 1.8: The undersigned hereby certifies that this paper or papers, as described herein, are being deposited in the United States Postal Service, as first class mail, in an envelope address to: Mail Stop Appeal Brief-Patents, Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450 on this __19__ day of July, 2004.

By Steven E. Dicke

Name: Steven E. Dicke

Appeal Brief to the Board of Patent Appeals and Interferences
of the United States Patent and Trademark Office
Appellant: Maria Castellanos et al.
Serial No.: 09/944,919
Filed: August 31, 2001
Docket No.: 10007912-1
Title: METHOD AND SYSTEM FOR MINING A DOCUMENT CONTAINING DIRTY TEXT

Appendix A

1.   (Previously Presented) A computer-implemented method for mining a document containing dirty text comprising:

removing an instance of dirty text within said document to produce a cleaned document having a content; and

performing a data mining operation on said cleaned document thereby deriving relevant information from said cleaned document and providing a summary of the content of said document.

2.   (Original) The method for mining a document containing dirty text as recited in Claim 1, wherein said removing further comprises replacing an instance of dirty text with a standard term.

3.   (Original) The method for mining a document containing dirty text as recited in Claim 1, wherein said removing further comprises removing an instance of computer code from said document.

4.   (Original) The method for mining a document containing dirty text as recited in Claim 1, wherein said removing further comprises removing a table from said document.

5.   (Original) The method for mining a document containing dirty text as recited in Claim 1, wherein said performing a data mining operation further comprises identifying a sentence within said cleaned document by identifying a beginning and an end of said sentence.

6.   (Original) The method for mining a document containing dirty text as recited in Claim 5, wherein said performing a data mining operation further comprises scoring and ranking said sentence.

7.   (Original) The method for mining a document containing dirty text as recited in Claim 6, wherein scoring said sentence further comprises:

Appeal Brief to the Board of Patent Appeals and Interferences
of the United States Patent and Trademark Office
Appellant: Maria Castellanos et al.
Serial No.: 09/944,919
Filed: August 31, 2001
Docket No.: 10007912-1
Title: METHOD AND SYSTEM FOR MINING A DOCUMENT CONTAINING DIRTY TEXT

selecting scoring techniques operable for summarizing non-narrative, grammatically incorrect text;

selecting scoring techniques operable for summarizing narrative, grammatically correct text; and

using said scoring techniques to score said sentence.

8.   (Original) The method for mining a document containing dirty text as recited in Claim 7, wherein said method further comprises generating a summary derived from said scored and ranked sentences.

9.   (Original) The method for mining a document containing dirty text as recited in Claim 1, wherein said method further comprises selecting a text mining component based upon said data mining operation to be performed.

10.   (Original) The method for mining a document containing dirty text as recited in Claim 1, wherein said method further comprises customizing said method by adjusting a parameter value.

11.   (Previously Presented) A computer system comprising:

a bus;

a memory unit coupled to said bus; and

a processor coupled to said bus, said processor for executing a method for mining a document containing dirty text comprising:

producing a cleaned document having a content comprising performing a general cleaning of said document by removing an instance of dirty text within said document including instances of misspelling and grammatical errors, and performing a domain and task specific cleaning of said document including removing instances of computer code and tables to produce a cleaned document; and

performing a data mining operation on said cleaned document including providing a summary of the content of said document.

Appeal Brief to the Board of Patent Appeals and Interferences
of the United States Patent and Trademark Office
Appellant: Maria Castellanos et al.
Serial No.: 09/944,919
Filed: August 31, 2001
Docket No.: 10007912-1
Title: METHOD AND SYSTEM FOR MINING A DOCUMENT CONTAINING DIRTY TEXT

12.   (Previously Presented) The computer system as recited in Claim 11, wherein said removing further comprises replacing an instance of dirty text with a standard term.

13.-14. (Cancelled)

15.   (Original) The computer system as recited in Claim 11, wherein said performing a data mining operation further comprises identifying a sentence within said cleaned document by identifying a beginning and an end of said sentence.

16.   (Original) The computer system as recited in Claim 15, wherein said performing a data mining operation further comprises scoring and ranking said sentence.

17.   (Original) The computer system as recited in Claim 16, wherein scoring said sentence further comprises:

selecting scoring techniques operable for summarizing non-narrative, grammatically incorrect text;

selecting scoring techniques operable for summarizing narrative, grammatically correct text; and

using said scoring techniques to score said sentence.

18.   (Previously Presented) The computer system as recited in Claim 17, wherein said method further comprises generating the summary derived from said scored and ranked sentences.

19.   (Original) The computer system as recited in Claim 11, wherein said method further comprises selecting a text mining component based upon said data mining operation to be performed.

20.   (Original) The computer system as recited in Claim 11, wherein said method further comprises customizing said method by adjusting a parameter value.

**Appeal Brief to the Board of Patent Appeals and Interferences
of the United States Patent and Trademark Office**
Appellant: Maria Castellanos et al.
Serial No.: 09/944,919
Filed: August 31, 2001
Docket No.: 10007912-1
Title: METHOD AND SYSTEM FOR MINING A DOCUMENT CONTAINING DIRTY TEXT

21. (Previously Presented) A computer-useable medium having computer-readable program code embodied therein for causing a computer system to perform the steps of:

removing an instance of dirty text within said document to produce a cleaned document having a content; and

performing a data mining operation on said cleaned document to provide a summary of said content.

22. (Original) The computer-useable medium of Claim 21, wherein said removing further comprises replacing an instance of dirty text with a standard term.

23. (Original) The computer-useable medium recited in Claim 21, wherein said removing further comprises removing an instance of computer code from said document.

24. (Original) The computer-useable medium recited in Claim 21, wherein said removing further comprises removing a table from said document.

25. (Original) The computer-useable medium recited in Claim 21, wherein said performing a data mining operation further comprises identifying a sentence within said cleaned document by identifying a beginning and an end of said sentence.

26. (Original) The computer-useable medium recited in Claim 25, wherein said performing a data mining operation further comprises scoring and ranking said sentence.

27. (Original) The computer-useable medium recited in Claim 26, wherein scoring said sentence further comprises:

selecting scoring techniques operable for summarizing non-narrative, grammatically incorrect text;

selecting scoring techniques operable for summarizing narrative, grammatically correct text; and

using said scoring techniques to score said sentence.

Appeal Brief to the Board of Patent Appeals and Interferences
of the United States Patent and Trademark Office
Appellant: Maria Castellanos et al.
Serial No.: 09/944,919
Filed: August 31, 2001
Docket No.: 10007912-1
Title: METHOD AND SYSTEM FOR MINING A DOCUMENT CONTAINING DIRTY TEXT

28.   (Original) The computer-useable medium recited in Claim 27, wherein said method further comprises generating a summary derived from said scored and ranked sentences.

29.   (Original) The computer-useable medium as recited in Claim 21, wherein said method further comprises selecting a text mining component based upon said data mining operation to be performed.

30.   (Original) The computer-useable medium as recited in Claim 21, wherein said method further comprises customizing said method by adjusting a parameter value.

31.   (Previously Presented) A computer-implemented method for mining a document containing dirty text comprising:

    producing a cleaned document having a content comprising performing a general cleaning of said document by removing one or more instance of dirty text within said document including instances of misspelling and grammatical errors, and performing a domain and task specific cleaning of said document including removing instances of computer code and tables; and

    performing a data mining operation on said cleaned document, including determining a sentence score for each sentence of said cleaned document and ranking the sentences from highest to lowest based on the sentence score;

    generating a summary of the content of the document using the highest ranked sentences.

32.   (Previously Presented) The method of claim 31, wherein determining a sentence score for each sentence includes applying a keyword technique to each sentence.

33.   (Previously Presented) The method of claim 32, wherein determining a sentence score further comprises applying a location technique to each sentence.

34.   (Previously Presented) The method of claim 32, wherein determining a sentence score further comprises applying a semantic similarity technique to each sentence.

Appeal Brief to the Board of Patent Appeals and Interferences
of the United States Patent and Trademark Office
Appellant: Maria Castellanos et al.
Serial No.: 09/944,919
Filed: August 31, 2001
Docket No.: 10007912-1
Title: METHOD AND SYSTEM FOR MINING A DOCUMENT CONTAINING DIRTY TEXT

35.   (Previously Presented) The method of claim 34, wherein the semantic similarity technique comprises:

generating a vector associated with each sentence; and

comparing each vector to every other vector, including defining a cosine of an angle between two vectors and using the cosine of the angle between two vectors to determine whether sentences represented by the two vectors are semantically related.